ELSEVIER

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT®

Journal of Computational and Applied Mathematics 156 (2003) 389–401

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Enclosing clusters of zeros of polynomials

Arnold Neumaier

Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Wien, Austria

Received 15 May 2001; received in revised form 3 December 2002

Abstract

Lagrange interpolation and partial fraction expansion can be used to derive a Gerschgorin-type theorem that gives simple and powerful a posteriori error bounds for the zeros of a polynomial if approximations to all zeros are available. Compared to bounds from a corresponding eigenvalue problem, a factor of at least two is gained.

The accuracy of the bounds is analyzed, and special attention is given to ensure that the bounds work well not only for single zeros but also for multiple zeros and clusters of close zeros.

A Rouché-type theorem is also given, that in many cases reduces the bound even further.

MSC: Primary 65H05; Secondary 65G10

Keywords: Polynomial zeros; Gerschgorin disk; Multiple roots; Root cluster; Eigenvalue problem; Rouché's theorem

1. Introduction

Polynomial zeros have a very long history, and a multitude of papers on the subject appeared. An exhaustive bibliography is available in [23], and a book in [26] is dedicated solely to this topic. Therefore, we only review work that was used during the research resulting in the present results.

Error bounds for zeros of polynomials usually take the form of disks in the complex plane $\mathbb{C}$, and circular arithmetic (see [13,17]) is frequently used to obtain such enclosing disks. While a lot of information is available on enclosures for isolated zeros (see, e.g., [7,27] for very accurate aposteriori enclosures), much less has been done on enclosing multiple zeros or tight cluster of zeros that are difficult to isolate numerically. Interesting results in this direction were obtained in [18,22].

E-mail address: neum@cma.univie.ac.at (A. Neumaier).
URL: http://solon.cma.univie.ac.at/~neum/

0377-0427/03/$ - see front matter
doi:10.1016/S0377-0427(03)00380-7

Most methods of the latter type are related to Gerschgorin's theorem. This theorem (see, e.g., [32]) states that all eigenvalues of a complex $n \times n$-matrix A belong to the union S of the closed disks $D[A_{jj}; r_j]$ with center A_{jj} and radius $r_j = \sum_{k \neq j} |A_{jk}|$. Moreover, every connected component of S consisting of m of these disks contains exactly m eigenvalues of A, if these are counted with their algebraic multiplicity.

Since the eigenvalues are the zeros of the characteristic polynomial of A, we expect that a similar result is possible for zeros of polynomials. Indeed, any polynomial $f(z)$ of degree n with highest coefficient 1,

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_n, \tag{1}$$

is the characteristic polynomial $f(z) = \det(zI - A)$ of the companion matrix

$$A = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

Applying Gerschgorin's theorem to A and A^{T} gives well-known but crude bounds for the zeros of the polynomial (1), cf. [3,34].

With suitable generalizations of companion matrices, Smith [31], Börsch-Supan [4], Braess & Hadeler [5], Elsner [12] and Carstensen [8,9] are able to use Gerschgorin's theorem to get good bounds for individual zeros and zero clusters when good approximations are available.

In the following, we obtain sharper Gerschgorin-type disks that provide elegant, cheap and quite accurate bounds for individual zeros of a polynomial if good approximations to all zeros are available. The polynomial need not be in standard form (1) but may have an arbitrary representation; it suffices that a program is available that computes function values (and for rigorous bounds enclosures of function values).

We also apply an idea of Neumaier [24] to the partial fraction expansion, and obtain a Rouché-type theorem that gives a simple sufficient condition that a disc contains as many zeros as approximate zeros. In many cases, this gives even better error bounds.

Many zerofinders [23] can be used to obtain crude or highly accurate approximations for all zeros. In this paper we are only concerned with bounding the errors in such approximations, and mostly ignore the question of how good approximations are computed in the first place. Also, after having obtained bounds, one can in principle use these as starting enclosures for various other methods (such as [30]) to squeeze out a little more accuracy.

Extensive tests of the new bounds, and comparisons to a number of other methods for enclosing multiple roots, were done in [29] with the INTLAB interval environment [28], with very positive results. Therefore, no numerical results are given here.

2. Partial fraction expansions

Let $f(z)$ be a polynomial of degree n with highest coefficient $a_0 \neq 0$, with (unknown) zeros $\zeta_1,\ldots,\zeta_n$ (repeated according to their algebraic multiplicity), so that

$$f(z) = a_0 \prod_{k=1}^{n} (z - \zeta_k). \tag{2}$$

Let $z_1,\ldots,z_n$ be (known) approximations to the zeros (with unknown accuracy), and define the polynomial

$$g(z) := \prod_{k=1}^{n} (z - z_k). \tag{3}$$

We assume that the z_k are distinct; then the partial fraction expansion of $f(z)/g(z)$ has the form

$$\frac{f(z)}{g(z)} = a_0 + \sum_{k=1}^{n} \frac{p_k}{z - z_k}. \tag{4}$$

The coefficients p_k in this expansion are easily identified as

$$p_k = f(z_k) \Big/ \prod_{l \neq k} (z_k - z_l) \tag{5}$$

(cf. [4]). To check this, it suffices to multiply (4) by $g(z)$ and observe that we simply obtain the Lagrange interpolation formula for interpolating the polynomial $f(z) - a_0 g(z)$ of degree $\leqslant n-1$ in the n points $z_1,\ldots,z_n$.

The p_k can be calculated easily whenever a program for evaluating the polynomial is available. Thus the polynomial need not be in standard form (1) but may have an arbitrary representation.

If the highest coefficient is not available, it can be computed by evaluating both sides of (4) for some $z \neq z_k$ and solving for a_0. In many applications, however, the highest coefficient a_0 is known; in particular, this holds for the case when the polynomial is given in standard form, as a Taylor series expansion, as a Newton interpolation polynomial. In a polynomial eigenvalue problem, where $f(z) = \det(A_0 z^s + \cdots + A_{s-1} z + A_s)$ with nonsingular $A_0 \in \mathbb{C}^{d\times d}$, the degree is $n = sd$ and the highest coefficient is $a_0 = \det A_0$. (The singular case is more tricky.) Note that in such cases, rigorous computations produce a small disk for a_0; it is not advisable to normalize the highest coefficient to 1 since this introduces unnecessary overestimation due to additional dependence.

Note that once a set of p_k is available one needs no further evaluations of $f(z)$ to determine the zeros of the latter; it suffices to find the zeros of the rational function

$$h(z) = a_0 + \sum_{k=1}^{n} \frac{p_k}{z - z_k}. \tag{6}$$

This is important when $f(z)$ is given implicitly by a lengthy program (such as in the polynomial eigenvalue problem mentioned earlier). Solving $h(z) = 0$ for $z - z_j$ gives the fixed point

equation

$$z = z_j - \frac{p_j}{a_0 + \sum_{k\neq j} p_k/(z - z_k)}. \tag{7}$$

Börsch-Supan [4] observed that if all p_k are small and z_j is not close to some other z_k then this defines a fixed point iteration converging quickly to the zero close to z_j; indeed, Nourein [25] proved local cubic convergence to isolated zeros. For clusters of zeros, the behavior is erratic but a spiral search [2] applied to $h(z) = 0$ will produce good approximations to the zeros in the cluster.

In particular, for expensive f, this can be used to compute refined representations of form (4) with improved approximations z_k' in place of z_k, without needing new evaluations of f. Indeed, the function values $f(z_k')$ needed in (5) can be found by evaluating instead (4) at $z = z_k'$ and multiplying with $g(z)$ from (3).

Zerofinding problems for functions of the form (6) also arise in Cuppen's [10] divide-and-conquer approach to the tridiagonal eigenvalue problem, and in the problem of updating the spectrum of a matrix under rank one changes, see [14,6,15].

3. A Gerschgorin-type theorem

We want to enclose the set of zeros in a union of complex disks. Writing z^* for the complex conjugate of z and $D[z;r] := \{z \in \mathbb{C} \mid z^*z \leqslant r^2\}$ for the closed disk with midpoint z and radius r, we have the following simple enclosure condition.

Lemma 3.1. *For real* $\alpha > 0$,

$$\mathrm{Re}\left(\alpha + \frac{a}{b}\right) \leqslant 0 \Rightarrow b \in D\left[-\frac{a}{2\alpha}; \left|\frac{a}{2\alpha}\right|\right].$$

Proof. From

$$\begin{aligned} 0 \geqslant b^*b\,\mathrm{Re}(\alpha + a/b) &= b^*b \cdot (\alpha + a/2b + a^*/2b) \\ &= \alpha b^*b + b^*a/2 + a^*b/2 = \alpha(|b + a/2\alpha|^2 - |a/2\alpha|^2), \end{aligned}$$

we obtain $|b + a/2\alpha| \leqslant |a/2\alpha|$, as claimed. □

This simple test has the following useful consequence.

Theorem 3.2. *If* $a_0 \neq 0$ *then all zeros of* $f(z)$ *belong to the union* S *of the disks*

$$D_j := D[z_j - np_j/2a_0; |np_j/2a_0|]. \tag{8}$$

Moreover, every connected component of S *consisting of* m *of these disks contains exactly* m *zeros of* $f(z)$, *if these are counted with their algebraic multiplicity.*

Proof. Substitution of an arbitrary zero $z=\zeta$ of $f(z)$ into (4) gives

$$0=\frac{n}{a_0}\operatorname{Re}\left(a_0+\sum_{k=1}^{n}\frac{p_k}{\zeta-z_k}\right)=\sum_{k=1}^{n}\operatorname{Re}\left(1+\frac{np_k}{a_0(\zeta-z_k)}\right).$$

This implies that for some index j,

$$\operatorname{Re}\left(1+\frac{np_j}{a_0(\zeta-z_j)}\right)\leqslant 0.$$

By Lemma 3.1, we conclude that $\zeta\in D[z_j-np_j/2a_0;|np_j/2a_0|]\subseteq S$. Since ζ was arbitrary, any zero of f is in S.

To prove the second part, note that for $0\leqslant t\leqslant 1$,

$$p_t(z):=a_0g(z)+\sum_{k=1}^{n}tp_k\frac{g(z)}{z-z_k}$$

defines a polynomial of degree n, and $p_0(z)=a_0g(z)$, $p_1(z)=f(z)$. For $t=0$, $p_t(z)$ therefore has the zeros z_k, and for $0\leqslant t\leqslant 1$, $p_t(z)$ has its zeros in the union of the disks

$$D[z_j-ntp_j/2a_0;ntp_j/2a_0]\subseteq S.$$

Since the zeros of a polynomial depend continuously on the coefficients, the m zeros of $p_0(z)$ in a component S_0 of S formed by m disks cannot leave S_0. Hence, all $p_t(z)$, $0\leqslant t\leqslant 1$, and in particular $f(z)=p_1(z)$ have m zeros in S_0. □

Remarks 3.3. (i) To apply the theorem we need approximations z_k for all roots. There are many zerofinders that can be used [23]; a common and efficient method is the Jenkins–Traub method [19,20]. We need to assume that the z_k are distinct even when $f(z)$ has multiple roots; this is easily achieved in practice, see the discussion in Section 4.

(ii) It is not difficult to show that $f(z)$ is the characteristic polynomial of the matrix

$$A=\operatorname{Diag}(z_k)-pe^{\mathrm{T}},$$

where $p=(p_1,\dots,p_n)^{\mathrm{T}}$, $e=(1,\dots,1)^{\mathrm{T}}$. Indeed, writing

$$u=(p_1/(z-z_1),\dots,p_n/(z-z_n))^{\mathrm{T}},$$

we have

$$\begin{aligned}\det(zI-A)&=\det(\operatorname{Diag}(z-z_k)+pe^{\mathrm{T}})=\det(\operatorname{Diag}(z-z_k)(I+ue^{\mathrm{T}}))\\&=\det\operatorname{Diag}(z-z_k)\det(I+ue^{\mathrm{T}})=\prod(z-z_k)(1+e^{\mathrm{T}}u)\\&=g(z)\left(1+\sum\frac{p_k}{z-z_k}\right)=f(z).\end{aligned}$$

Thus we may apply Gerschgorin's theorem to this matrix (cf. [31,4,5,12]) and find that the eigenvalues of A (i.e., the zeros of $f(z)$) belong to the union of the disks $D[z_j-p_j;(n-1)p_j]$. These disks contain the disks of our theorem but have (for large n) almost twice the diameter. Thus

Theorem 3.2 is significantly stronger than Gerschgorin's theorem for that matrix, unless the approximations to the zeros are so good that the disks computed as an enclosure of $np_j/2a_0$ are nearly centered at zero.

4. Accuracy of the bounds

The quotients p_k/a_0 are the corrections to the approximate zeros z_k in the (for simple zeros quadratically convergent) Durand–Kerner method [11,21] that actually goes back to Weierstrass [33] (see [16]; I learnt it from the recent Ph.D. thesis in [1]). Thus the bounds are just a factor $n/2$ larger than these corrections.

The same factor is obtained from an analysis of the attainable limit accuracy. Even the best numerical method will produce approximations z_k such that the corresponding polynomial $g(z) = \prod(z - z_k)$ only satisfies $f(z) - a_0 g(z) = \mathrm{O}(\varepsilon)$, where ε is the machine precision. Thus let us assume first that

$$f(z) - a_0 g(z) = \varepsilon d(z) \tag{9}$$

with a polynomial $d(z)$ of degree $< n$ and coefficients of order O(1). Standard perturbation theory (see, e.g., [32, Chapter 5.8]) implies that the zeros of $f(z)$ have the form

$$\zeta_k = z_k - \varepsilon \frac{d(z_k)}{a_0 g'(z_k)} + \mathrm{O}(\varepsilon^2).$$

Since $f(z_k) = a_0 g(z_k) + \varepsilon d(z_k) = \varepsilon d(z_k)$ and

$$g'(z_k) = \sum_j \prod_{l \neq j} (z - z_j)|_{z=z_k} = \prod_{l \neq k} (z_k - z_l),$$

we see that, unless z_k is very close to some other z_l,

$$p_k = \varepsilon \frac{d(z_k)}{g'(z_k)} = \mathrm{O}(\varepsilon)$$

and

$$\zeta_k - z_k = p_k/a_0 + \mathrm{O}(\varepsilon^2).$$

Thus the radius of our disks is about a factor $n/2$ larger than the error in the approximations if the linearized perturbation theory is applicable. This is the case whenever z_k is close to a simple and well isolated zero of $f(z)$, and this shows numerically in that the disk $D[z_k - np_k/2a_0; np_k/2a_0]$ is a connected component of S.

4.1. Multiple zeros and zero clusters

Let C be the cluster of m closest zero approximations z_k $(k = 1, \ldots, m)$ to a given center z_C; we take $z_C := m^{-1} \sum_{k \leqslant m} z_k$ if the cluster is given. To improve the quality of the enclosure we may

replace the approximations z_k $(k = 1,\dots,m)$ in each cluster C by better separated approximations

$$z_k := z_C + \rho_k s \tag{10}$$

with a suitable (cluster dependent) $s \neq 0$ and the mth roots of unity ρ_k. With this new approximations, we define

$$q_C(z) = \frac{f(z)}{\prod_{z_k \notin C}(z - z_k)}, \tag{11}$$

$$\varepsilon_C = \max_{z_k \in C} |\tilde{q}_C(z_k) - \tilde{q}_C(z_C)|, \tag{12}$$

where the tilde denotes computed values. Then

$$g_C(z) := \prod_{k \leqslant m} (z - z_k) = (z - z_C)^m - s^m,$$

$$\prod_{l \leqslant m,\ l \neq k} (z_k - z_l) = g_C'(z_k) = m(z_k - z_C)^{m-1} = \rho_k^{-1} m s^{m-1},$$

hence

$$p_k = \frac{\rho_k q_C(z_k)}{m s^{m-1}}. \tag{13}$$

If we evaluate (13) in finite precision arithmetic, we expect

$$|\tilde{q}_C(z_k) - q_C(z_k)| \approx \delta := \frac{\delta_f}{\prod_{z_k \notin C} |z_C - z_k|},$$

where δ_f is an estimate for the rounding error obtainable by a running error analysis [35] or in circular arithmetic (see [17]) as rad $f(\mathbf{z}_C)$, where $\mathbf{z}_C$ is a very narrow interval around z_C.

Since the $z_k \notin C$ are approximate zeros, we also expect

$$q_C(z_k) \approx a_0 \prod_{l \leqslant m} (z_l - \zeta_l) \approx a_0 (z_k - z_C)^m = a_0 s^m,$$

hence, the computed value $\tilde{p}_k$ satisfies

$$|\tilde{p}_k| \approx \frac{|a_0 s^m| + \delta}{|m s^{m-1}|} = \frac{1}{m}(|a_0 s| + \delta |s|^{1-m}).$$

The right hand side is minimal for

$$|s| = \sigma := \left(\frac{(m-1)\delta}{|a_0|}\right)^{1/m}, \tag{14}$$

and for this choice,

$$|\tilde{p}_k| \approx \frac{1}{m}(|a_0|\sigma + \delta\sigma^{1-m}) = \frac{|a_0|\sigma}{m-1}$$

has the magnitude of the limit accuracy for multiple zeros. The phase of s may still be chosen freely, and a natural choice is to take

$$s = \frac{\sigma\tilde{s}}{|\tilde{s}|} \quad \text{where } \tilde{s} = \sqrt[m]{\prod_{z_k \in C} (z_k - z_C)}. \tag{15}$$

Thus calculation according to (11)–(15) should provide nearly optimal results even for multiple zeros. But in fact, zerofinders that find all roots simultaneously usually produce automatically approximations roughly separated as in (10). Thus, in practice, no special precautions are needed, and the above analysis applies to the original approximations z_j.

5. Refinement of the bounds

Once one has a partition of the zero set into clusters belonging to different connected component of S, it is possible to refine the enclosure (and possibly the partition) by applying the method used in the proof to each cluster separately. Indeed, denote by C_ν the set of indices j such that z_j is in the νth cluster, and by $|C_\nu|$ its size. Then the sets C_ν partition the index set $\{1,\dots,n\}$, and the cluster sums

$$r_\nu(z) := \frac{1}{a_0} \sum_{j \in C_\nu} \frac{p_j}{z - z_j}$$

satisfy

$$\sum_\nu r_\nu(\zeta) = -1 \quad \text{for all zeros } \zeta. \tag{16}$$

From the current enclosure we can compute circular enclosures $r_\mu(D_j)$ for $\{r_\mu(z) \,|\, z \in D_j\}$, where D_j is given by (8), and hence finite lower bounds λ_ν with

$$\lambda_\nu \leqslant \sum_{\mu \neq \nu} \operatorname{Re} r_\mu(z) \quad \text{for all } z \in D_j,\ j \in C_\nu. \tag{17}$$

Theorem 5.1. *If*

$$\alpha_\nu := \frac{(1 + \lambda_\nu)|a_0|}{|C_\nu|} > \frac{a_0}{n}, \tag{18}$$

then the zeros in the νth cluster belong to the union of the disks

$$D'_j := D\left[z_j - \frac{p_j}{2\alpha_\nu}; \left|\frac{p_j}{2\alpha_\nu}\right|\right] \tag{19}$$

with $j \in C_\nu$. Moreover, any connected component of this union consisting of m such disks contains m zeros of $f(z)$.

Proof. Let ζ be a zero in the vth cluster. Then ζ is in some D_k ($k \in C_v$), and we find from (16) and (17) that $\operatorname{Re} r_v(\zeta) \leqslant -1 - \lambda_v$. Using (18), we can rewrite this in the form

$$\sum_{j \in C_v} \left(\operatorname{Re} \frac{p_j}{\zeta - z_j} + \alpha_v \right) \leqslant 0,$$

so that, for some index $j \in C_v$,

$$\operatorname{Re} \frac{p_j}{\zeta - z_j} + \alpha_v \leqslant 0. \tag{20}$$

Now Lemma 3.1 implies that the disk (19) contains ζ. In particular, ζ lies in the union of the disks (19) with $j \in C_v$. Since ζ was an arbitrary zero in the cluster, the first part follows. The second part follows as in Theorem 3.2 by a continuation argument. □

Note that the radius in Theorem 3.2 has the form $n/2 \cdot |p_j/a_0|$, while the new radius is $|C_v|/(2+2\lambda_v) \cdot |p_j/2a_0|$. In practice, λ_v is usually close to zero. Then (18) holds, and the new radius overestimates the distance of ζ from the center (which in the linear approximation is approximately $|p_j/2a_0|$) only by a factor of about $\frac{1}{2}|C_v|$.

Whenever (18) holds, the radius of all D'_j (from the cluster) are strictly smaller than that of the corresponding D_j; The zeros in the vth cluster now belong to the union of the disks (19) with $j \in C_v$, and if this set has several components we may again split clusters. We may repeat the procedure until no further gain is possible. To have a more compact (but less accurate) final representation, one can compute for each final connected component of disks $D_j = D[\tilde{z}_j; \tilde{r}_j]$ ($j \in C_v$) the common midpoint and radius

$$\bar{z}_v = |C_v|^{-1} \sum_{j \in C_v} \tilde{z}_j, \qquad \bar{r}_v := \max_{j \in C_v} (\tilde{r}_j + |\tilde{z}_j - \bar{z}_v|), \tag{21}$$

to get a single disk $\bar{D}_v = D[\bar{z}_v; \bar{r}_v]$ enclosing the corresponding cluster of zeros. The union of these disks covers all zeros, and although overlap is possible, *each* disk $\bar{D}_v$ contains at least $|C_v|$ zeros.

Algorithm 5.2. *Step* 1: Find the connected components of $S = \bigcup D_j$ and put the corresponding (index sets of) clusters into a list of initial clusters.

Step 2: While the list of initial clusters is nonempty: Pick one of these clusters C_v and compute λ_v and α_v.

- If (18) holds, replace the disks $\bar{D}_j$ ($j \in C_v$) by D'_j from (19), find the connected components of *their* union, and insert them in place of C_v into the list of initial clusters.
- But if (18) does not hold, move C_v from the list of initial clusters to the list of final clusters.

Step 3: Compute for each cluster C_v in the list of final clusters the disk $\bar{D}_v = D[\bar{z}_v; \bar{r}_v]$ and the number $n_v = |C_v|$ of zeros in $\bar{D}_v$.

If this is done in circular arithmetic, and we have rigorous enclosures for the p_k defined by (5) then all bounds are rigorous and we have verified enclosures for all zeros of the polynomial. Of

course, the quality of the bounds depends on the quality of the approximate zeros used, and improve with the accuracy of the approximations.

The partition into clusters found by the algorithm is of course not necessarily the same as that found by other methods; this is unavoidable since there is no natural unambiguous criterion for discriminating between one larger cluster and several smaller ones, or a cluster and isolated zeros.

6. A Rouché-type theorem

Lemma 6.1. *If* $z \in \partial D[c;r]$, $z_0 \notin \partial D[c;r]$, *then*

$$\left|\frac{p}{z-z_0} - \frac{(c-z_0)^* p}{|c-z_0|^2 - r^2}\right| = r\left|\frac{p}{|c-z_0|^2 - r^2}\right|. \tag{22}$$

Proof. For $|c-z_0| < r$, the same inequality follows from circular arithmetic; however, we need it for $|c-z_0| > r$. Let

$$q = \frac{c-z_0}{z-z_0}, \qquad s = \frac{r}{|c-z_0|}. \tag{23}$$

From

$$|q-1| = \left|\frac{c-z}{z-z_0}\right| = \frac{r}{|z-z_0|} = s|q|,$$

we find

$$\left|q - \frac{1}{1-s^2}\right|^2 = q^*q - \frac{q+q^*}{1-s^2} + \frac{1}{(1-s^2)^2}$$

$$= \frac{(q-1)^2 - s^2|q|^2}{1-s^2} + \frac{s^2}{(1-s^2)^2} = \frac{s^2}{(1-s^2)^2},$$

whence

$$\left|q - \frac{1}{1-s^2}\right| = \left|\frac{s}{1-s^2}\right|.$$

If we substitute (23) we find

$$\left|\frac{c-z_0}{z-z_0} - \frac{|c-z_0|^2}{|c-z_0|^2 - r^2}\right| = r\left|\frac{c-z_0}{|c-z_0|^2 - r^2}\right|,$$

and (22) follows upon multiplication with $p/(c-z_0)$. □

The following theorem is inspired by applications of Rouché's theorem (see, e.g., [17]) in [4], but it uses a slight sharpening of Rouché's theorem used implicitly in [24] and explicitly formulated by Batra [1].

Theorem 6.2. *If*

$$\mathrm{Re}\left(a_0+\sum_k \frac{(c-z_k)^* p_k}{|c-z_k|^2-r^2}\right) > r\sum_k \left|\frac{p_k}{|c-z_k|^2-r^2}\right| \tag{24}$$

then f and g have the same number of zeros in $D[c;r]$.

Proof. For $z\in\partial D[c;r]$,

$$\begin{aligned}
&\left|\mathrm{Re}\,\frac{f(z)}{g(z)}-\mathrm{Re}\left(a_0+\sum_k \frac{(c-z_k)^* p_k}{|c-z_k|^2-r^2}\right)\right| \\
&\quad=\left|\sum_k \mathrm{Re}\left(\frac{p_k}{z-z_k}-\frac{(c-z_k)^* p_k}{|c-z_k|^2-r^2}\right)\right| \\
&\quad\leqslant \sum_k \left|\frac{p_k}{z-z_k}-\frac{(c-z_k)^* p_k}{|c-z_k|^2-r^2}\right| \\
&\quad=\sum_k r\left|\frac{p_k}{|c-z_k|^2-r^2}\right|
\end{aligned}$$

by Lemma 6.1. The assumption of the theorem, therefore, implies that

$$\mathrm{Re}\,\frac{f(z)}{g(z)}>0 \quad \text{for all } z\in\partial D[c;r].$$

Hence $\mathrm{Re}(f(z)/g(z))$ has a constant sign on the boundary of $D[c;r]$. The principle of the argument [17] now implies that the number of zeros and the number of poles of f/g in $D[c;r]$ are the same. (Rouché's theorem asserts the same statement under the stronger condition $|f(z)|>|f(z)-g(z)|$ on $\partial D[c;r]$.) But the zeros of f/g are the zeros of f and the poles of f/g are the zeros of g. □

One may search for the best enclosing discs by varying c and r, minimizing r subject to the constraint (24). This may be especially profitable for the enclosure of root clusters, where fixed choices of c and r seem vulnerable to the presence of an additional zero not belonging to the cluster.

The special case

$$a_0=1, \qquad c=z_1, \qquad 0<r<r_{\max}=\min_{k>1}|c-z_k|$$

is instructive. The condition of the theorem reduces to

$$\varphi_0(r)>\varphi(r),$$

where

$$\varphi_0(r) = \mathrm{Re}\left(1 + \sum_{k>1} \frac{(c - z_k)^* p_k}{|c - z_k|^2 - r^2}\right),$$

$$\varphi(r) = \frac{|p_1|}{r} + r \sum_{k>1} \left| \frac{p_k}{|c - z_k|^2 - r^2} \right|.$$

If all $p_k = \mathrm{O}(\varepsilon)$, $\varepsilon \ll r_{\max}$ then the choice $r := \kappa |p_1|$ with fixed $\kappa > 1$ gives

$$r = \mathrm{O}(\varepsilon), \qquad \varphi_0(r) = 1 + \mathrm{O}(\varepsilon), \qquad \varphi(r) = \frac{1}{\kappa} + \mathrm{O}(\varepsilon^2).$$

Thus if $\varepsilon \ll \kappa - 1$, Theorem 6.2 applies and shows the existence of a unique zero in $D[z_1; \kappa |p_1|]$. Thus the disks computable from Theorem 6.2 do not involve the spurious factor $n/2$ that occurs in the Gerschgorin-type theorem.

For isolated zeros, a similar result (also without the spurious factor $n/2$) was proved in [27]. However, since the latter involves a bound on the separation of all zeros, *all* enclosures suffer from the presence of a *single* pair of close zeros.

Acknowledgements

I want to thank Siegfried Rump for inspiring discussions about enclosing clusters, and the referees for many comments that lead to an improved presentation of the material.

References

[1] P. Batra, Abschätzungen und Iterationsverfahren für Polynom-Nullstellen, Dissertation, Technische Univ. Hamburg-Harburg, 1999.
[2] F. Bauhuber, Direkte Verfahren zur Berechnung der Nullstellen von Polynomen, Computing 5 (1970) 97–118.
[3] H.E. Bell, Gershgorin's theorem and the zeros of polynomials, Amer. Math. Monthly 72 (1965) 292–295.
[4] W. Börsch-Supan, Residuenabschätzungen für Polynom-Nullstellen mittels Lagrange-Interpolation, Numer. Math. 14 (1970) 287–296.
[5] D. Braess, K.P. Hadeler, Simultaneous inclusion of the zeros of a polynomial, Numer. Math. 21 (1973) 161–165.
[6] J.R. Bunch, C.P. Nielsen, D.C. Sorensen, Rank one modification of the symmetric eigenproblem, Numer. Math. 31 (1978) 31–48.
[7] C. Carstensen, Anwendungen von Begleitmatrizen, Z. Angew. Math. Mech. 71 (1991) T809–T812.
[8] C. Carstensen, Inclusion of the roots of a polynomial based on Gershgorin's theorem, Numer. Math. 59 (1991) 349–360.
[9] C. Carstensen, E. Stein, Über die Falksche ECP-transformation und Verallgemeinerungen, Z. Angew. Math. Mech. 69 (1989) 375–391.
[10] J.J.M. Cuppen, A divide and conquer method for the symmetric tridiagonal eigenproblem, Numer. Math. 36 (1981) 177–195.
[11] I.E. Durand, Solutions Numérique des Équations Algébriques. Tome I: Équations du Type F(x) = 0, Racines d'une Polynôme Masson, Paris, 1960, pp. 279–281.
[12] L. Elsner, A remark on simultaneous inclusions of the zeros of a polynomial by Gershgorin's theorem, Numer. Math. 21 (1973) 425–427.

[13] I. Gargantini, P. Henrici, Circular arithmetic and the determination of polynomial zeros, Numer. Math. 18 (1972) 305–320.
[14] G.H. Golub, Some modified matrix eigenvalue problems, SIAM Rev. 15 (1973) 318–344.
[15] M. Gu, S.C. Eisenstat, Stable and efficient algorithm for the rank one modification of the symmetric eigenproblem, SIAM J. Matrix Anal. Appl. 15 (1994) 1266–1276.
[16] P. Henrici, Uniformly convergent algorithms for the simultaneous determination of all zeros of a polynomial, in: J.M. Ortega, W.C. Rheinboldt (Eds.), Proceedings of the Symposium on Numerical Problems, Studies in Numerical Analysis, Vol. 2, SIAM, Philadelphia, 1970, pp. 1–8.
[17] P. Henrici, Applied and Computational Complex Analysis, Vol. 1, Power Series—Integration—Conformal Mapping—Location of Zeros, Wiley, New York, 1974.
[18] V. Hribernig, H.J. Stetter, Detection and validation of clusters of polynomial zeros, J. Symbolic Comput. 24 (1997) 667–681.
[19] M.A. Jenkins, Algorithm 493: zeros of a real polynomial, ACM Trans. Math. Software 1 (1975) 178–189.
[20] M.A. Jenkins, J.F. Traub, A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration, Numer. Math. 14 (1970) 252–263.
[21] I.O. Kerner, Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, Numer. Math. 8 (1966) 290–294.
[22] P. Kravanja, T. Sakurai, M. Van Barel, On locating clusters of zeros of analytic functions, BIT 39 (1999) 646–682.
[23] J.M. McNamee, A bibliography on roots of polynomials, J. Comput. Appl. Math. 47 (1993) 391–394, Available online at http://www.elsevier.com/homepage/sac/cam/mcnamee.
[24] A. Neumaier, An existence test for root clusters and multiple roots, Z. Angew. Math. Mech. 68 (1988) 256–257.
[25] A.W. Nourein, An iteration formula for simultaneous determination of the zeroes of a polynomial, J. Comput. Appl. Math. 1 (1975) 251–254.
[26] M.S. Petkovic, Iterative Methods for Simultaneous Inclusion of Polynomial Zeros, Lecture Notes in Mathematics, Vol. 1387, Springer, Berlin, 1989.
[27] M.S. Petkovic, D. Herceg, S. Ilic, Point estimation and some application to iterative methods, BIT 38 (1998) 111–126.
[28] S.M. Rump, INTLAB—interval laboratory, in: T. Csendes (Ed.), Developments in reliable computing, Kluwer, Dordrecht, 1999, pp. 77–104, http://www.ti3.tu-harburg.de/~rump/intlab/index.html.
[29] S.M. Rump, Ten methods to bound multiple roots of polynomials, J. Comput. Appl. Math., this issue.
[30] T. Sakurai, H. Sugiura, Improvement of convergence of an iterative method for finding polynomial factors of analytic functions, J. Comput. Appl. Math. 140 (2002) 713–725.
[31] B.T. Smith, Error bounds for zeros of a polynomial based upon Gerschgorin's theorems, J. ACM 17 (1970) 661–674.
[32] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, Berlin, 1987.
[33] K. Weierstrass, Neuer Beweis des Satzes, daß jede ganze rationale Function einer Veränderlichen dargestellt werden kann als ein Product aus linearen Functionen derselben Veränderlichen, in: Ges. Math. Werke 3, Meyer und Muller, Berlin, 1903, pp. 251–269 (reprinted by Johnson Reprint Corporation, New York, 1967.).
[34] W. Werner, Über Abschätzungen von Polynomnullstellen mittels des Gerschgorinschen Kreissatzes, Z. Angew. Math. Mech. 63 (1983) T390–T391.
[35] J.H. Wilkinson, Rounding Errors in Algebraic Processes, Prentice-Hall, Englewood Cliffs, NJ, 1963.